

TS-EAS EAC-CPF Berlin meeting

Attendees (in person)

Wednesday, 11 March 2020

EAD3 reconciliation

Namespace

Language related attributes

Related standards via @encodinganalog

Formatting with @altrender and @label

Dates

Abbreviated names

Mixed content in the standards

<control> and its sub-elements as case study for shared schema option

<control> element in itself

<recordId> and <otherrecordId>

<maintenancestatus>

<publicationstatus>

<maintenanceagency>

Wrap-up and summary

Decisions

Attendees (in person)

Name	Institution	Role
Cory Nimer	Brigham Young University, US	Documentation team
Gerhard Müller	Staatsbibliothek zu Berlin, DE	TS-EAS ex officio
Joost van Koutrik	Het Utrechts Archief, NL	TS-EAS
Karin Bredenberg	Kommunalförbundet Sydarkivera, SE	TS-EAS co-chair
Kerstin Arnold	Archives Portal Europe Foundation, EU	EAD team lead
Regine Heberlein	Princeton University, US	former EAC-CPF team member
Silke Jagodzinski	Geheimes Staatsarchiv Preußischer Kulturbesitz, DE	EAC-CPF team lead

Wednesday, 11 March 2020

EAD3 reconciliation

Kerstin is going to introduce today's topic: EAD3 Reconciliation. Caitlin and Kerstin prepared a paper on shared and similar elements and attributes in EAD3 and EAC-CPF. The aim is to ensure that shared elements, i.e. elements that use the same element names, really are of the same scope and are hence using shared definitions in both standards. Similar elements on the other hand must not use the same name and need to be defined differently.

Namespace

We already agreed to remove xlink namespace from EAC-CPF.

Keeping or removing XML namespace is not decided yet. It was introduced with EAC-CPF, i.e. EAC beta did not have it, nor did any EAD version to date. Schema team is asked to dive into this topic:

- look for examples of other schemas that removed the XML namespace or are using their own variation of XML namespace attributes (as e.g. METS does),
- look at the attributes definition to decide, if keeping the XML namespace spares us to define the attributes in questions (especially @xml:lang) within the EAS or related Schematron.

A decision should be made until our next in-person meeting in Chicago.

Language related attributes

The set of base attributes in EAD3 includes the attribute [@lang](#) that is in fact available in all elements, except [<ptr>](#). To keep in mind for the usage of EAC-CPF decision on @(xml:)lang attribute.

The EAC-CPF Tag Library defines that @xml:lang contains values in a list taken from the IANA Registry. This list was never defined in the schema.

Decision: Language codes should be added to the schematron.

The attribute [@transliteration](#) is under discussion to keep it at all or to remove. As we have use cases in the EAC-CPF team, we agreed to keep the attribute. We could even find use cases to use different transliterations in one file (e.g. having aggregate instances or having descriptions with transliterations from different languages). Rules for transliteration should be declared via conventionDeclaration.

Decision: Keep @translation attribute as optional attribute along @(xml:)lang and @script in all non-empty elements in EAC-CPF.

Pending question: Need to decide what “non-empty” means and if there could be cases of “empty” elements, where it still makes sense to allow for language attribution, e.g. plural elements. There would ideally be the same approach taken for EAC-CPF and EAD.

Related standards via @encodinganalog

EAD3 used the attributes @relatedencoding and @encodinganalog for cross-references with related standards. @relatedencoding is available for <ead>, <control> and <archdesc>

to declare the standard referred to. @encodinganalog is then used on the single elements that have an equivalent in that standard.

At the moment, however, @encodinganalog is only available with elements that are mentioned in the crosswalks with ISAD(G) and MARC. This could be more open by adding @encodinganalog to all elements, leaving users the choice to use any standard to refer to, not only ISAAR(CPF) or ISAD(G) or MARC.

Decision: Add @encodinganalog as optional attribute with NMTOKEN in all elements in EAC-CPF.

Formatting with @altrender and @label

We already agreed to not to add both formatting attributes to EAC-CPF.

- A use case for [@altrender](#) in EAD3 is indexing titles within EAD3 at Princeton University. The noun marker in a title beginning would be marked with an @altrender.
- Kalliope uses [@label](#) for facetting information within EAD.

Decision: The EAD team will review this issue in one of their next meetings.

Dates

Some discussion about the encoding of date information, whether it was possible to implement EDTF (listed as [draft](#) on the Library of Congress website) and its status relative to the ISO 8601:2019 revisions. Additional information is needed before knowing how to proceed.

Abbreviated names

According to policies and usages for programmers it is recommended to use descriptive and non-abbreviated names.

As we already agreed to remove the camelCase in spelling the names within the schema, we agree to change the abbreviated names in EAC-CPF into descriptive names also for the already existing elements in that context.

Decision: Change abbreviated names in EAC-CPF into descriptive names.

Mixed content in the standards

EAD3 allows mixed content to various extent and on multiple elements. In the context of the EAD3 reconciliation, this specifically refers to the shared elements <abstract>, <addressline>, <citation>, <date>, <fromdate>, <todate>, <event>, <item>, <p>, and <part>.

Discussion is going around the need and usage of mixed content in the standards. Are the standards meant to encode narrative text created by ambitious archivists.

<control> and its sub-elements as case study for shared schema option

Originally developed in EAC-CPF and incorporated in EAD during the development of EAD3. Includes circa ⅓ of shared elements, making it a good element for reconciliation.

Furthermore it covers quite a few general questions when looking at its sub-elements in more detail.

<control> element in itself

Main difference is the continued existence of <filedesc> in EAD3, and currently EAD3 also includes <representation>. There already exists the suggestion that add <representation> to EAC-CPF (see [#9](#)), but what to do with <filedesc>?

Some sense that there really isn't a great need for <filedesc> in EAD3 either - title is not significantly different from finding aid title, author is most likely the processor, the publisher is probably the publisher of the finding aid. There does not appear to be a significant need for <filedesc> in EAC-CPF. TS-EAS could consider moving it out of <control> in EAD3.

Alternatively, could have it included as optional in the shared schema and deal with difference when publishing the respective schema files for EAC-CPF and EAD3.

<representation> in EAD3 provides a link to a transformed/deliverable version of the EAS instance. It is an optional, repeatable element with a range of attributes, including: @xml:id, @xml:lang, @script, @transliteration, @encodinganalog, and @localtype (this is one of several examples of @localtype in <control> that needs to be remediated). Pending question with regard to the linking attributes (previously XLink): which ones should be kept?

Similar to <filedesc>, <representation> does not seem to have a strong use case for EAC-CPF, but could be included as an optional element in the shared schema. As an example for a use case in EAD: in the context of Archives Portal Europe many institutions host their own finding aids, but provide an EAD file to the service for aggregation. The link to the original institution's finding aid would then provided in the <representation> element. A similar example for EAC-CPF could be coming from SNAC (e.g.

<https://snaccooperative.org/view/44995506>, where the EAC-CPF document currently does not point back directly to this web-based representation).

Decision: Add the element <representation> as it is used and defined in EAD3 for EAC-CPF within <control>.

On the question of @localtype, the expected use with <representation> would be for specifying the format or type of representation. However, this could potentially be specified in @linkrole instead. Could be part of the more general conversation about @localtype vs context-specific @...type attributes vs simply @type.

If we are moving toward using more specific attribute values, then should we provide value lists for the use of these fields? If so, how do we deal with outliers? Also, do we actually have a list of controlled values that we want to enforce? For example, to replace @localtype in <agencycode>, then we would need an @agencycodetype (for a controlled or uncontrolled list) and an @otheragencycodetype (to cover outliers if @agencycodetype is controlled). Which approach to take (particularly whether @other...type is needed) would vary based on the element involved. Looking at a potential <agencycode> @agencycodetype again as an example, it was suggested that even having a short list of values (isil or other) could still be valuable in analyzing data.

Suggested that we move to more specific attribute values in place of @localtype, but not have them controlled. This can then be addressed in the tag library with recommendations for values, but not enforcing them via the schema. However, there was no agreement that @localtype needs to be removed.

It was pointed out that @localtype, while @localtype can be used with any content, the expectation for the use of @localtype has been that locally used value lists would be defined in the <localtypedeclaration> section within <control>.

Need to refer this back to the Schema Team to determine whether it is a design principle to not include @localtype in <control> elements. Also need some more examples to work with for different elements, which might need more specific attributes. Also need to send along to the Schema Team the option of reinstating @type element that is specific to each element, or whether the use of element-specific type attributes is a design decision for the standard. Should be possible to resolve at the May meeting of the Schema Team. Depending on this decision, the question of specific named attributes should be tackled afterwards.

Decision: Hand a decision over to the Schema team, if @localType can be used in <control> elements.

Decision: Hand a discussion on reintroducing @type over to the TS-EAS.

Beyond @localtype, the EAD3 model currently includes a range of former XLink attributes for many elements. Proposed to retain @linkrole, @linktitle, and @href, and drop the others. Will wait to see if there are concerns from the community during the standard review.

Decision: Use @linkrole, @linktitle, and @href, from former XLink attributes and drop @actuate, @arcrole, @show.

<recordId> and <otherrecordId>

Suggestion that we widen the [<otherrecordId>](#) to also capture identifiers of instances aggregated with the current one ([#54](#)).

Some questions arose surrounding the original decision to record agency IDs as a separate element, if they could be handled through an attribute instead. One possible reason for keeping it an element would be if you needed to provide attributes about/defining them, as in our @agencycodetype example.

Suggested that @valueURI, @vocabularysource, and @vocabularysourceURI be added, as optional attributes, to encode the provenance of other identifiers. Could also use MODS's @authority and @authorityURI - but preferable to reuse existing terms in our own standards, seeing that @vocabularysource attribute already exists in the EAC-CPF format.

For example:

```
<otherrecordid vocabularysource="[code value for repository]"  
vocabularysourceURI="[original instance repository URI]" valueURI="[URI of  
original instance]">
```

Currently [<recordid>](#) is text in EAD3, but specifically NMTOKEN in EAC-CPF. Felt that there was no reason for these to be different, with a preference for moving this toward NMTOKEN. Will need to consult with the EAD Team about addressing this difference.

Decision: Discuss difference btw content resp. data type in <recordId> with EAD team.

<maintenancestatus>

Reviewed the availability of the element, asking whether it should be allowed to be repeatable with language, or whether the @(xml:)lang should be removed and the element should only be used with the list of @value attributes.

Decision: Use <maintenancestatus> as an empty element without text or other content with a required attribute @value with limited values and without language attributes. Seeing that the element is part of control and not part of the descriptive section of EAC-CPF, the

necessity to have language attribution and an option for internationalisation does not seem to be that prominent. The element is mandatory and not repeatable.

<publicationstatus>

Determined to follow the same course as above, and remove the @(xml:)lang attribute and require the use of the controlled list of @value terms.

Decision: Use <publicationstatus> as an empty element without text or other content with a required attribute @value with limited values and without language attributes. Seeing that the element is part of control and not part of the descriptive section of EAC-CPF, the necessity to have language attribution and an option for internationalisation does not seem to be that prominent. The element is optional and not repeatable.

<maintenanceagency>

Element is meant to note the institution or service responsible for EAS instance. It was suggested that we widen the scope and re-use either the element itself or its sub-elements <agencyname> and <agencycode> within <conventiondeclaration> ([#67](#)) in order to state who is responsible for the maintenance of the convention. Preference would be with using the sub-elements rather than the entire <maintenanceagency> wrapper).

Ultimately, it was decided that due to lack of follow-up from the original ticket creator, it would be preferable to close the ticket without action rather than making this change.

Decision: No additional agency information in convention declaration as there is no use case available.

Wrap-up and summary

Daily summary for TS-EAS members in a virtual session.

Following wrap-up, questions came up about whether it would be possible for online participants to add questions or comments to the meeting minutes.

It was decided that the preference would be to have others comment on the issues in GitHub if they have other opinions, rather than adding comments to the minutes. The minutes should be seen just as a record of the discussion at the in-person meeting.

Decisions

Topic	Decision	Issue
Language encoding	Language codes should be added to the schematron. Keep @translation attribute as optional attribute along @(xml:)lang and @script in all non-empty elements in EAC-CPF.	#28
EAD3 Reconciliation	Add @encodinganalog as optional attribute with NMToken in all elements in EAC-CPF.	#123

	Add the element <representation> as it is used and defined in EAD3 for EAC-CPF within <control>.	#81
	<p>Hand a decision over to the Schema team, if @localType can be used in <control> elements.</p> <p>Hand a discussion on reintroducing @type over to the TS-EAS.</p>	
	Use @linkrole, @linktitle, and @href, from former XLink attributes and drop @actuate, @arcrole, @show.	
	Use <maintenancestatus> as an empty element without text or other content with a required attribute @value with limited values and without language attributes. Seeing that the element is part of control and not part of the descriptive section of EAC-CPF, the necessity to have language attribution and an option for internationalisation does not seem to be that prominent. The element is mandatory and not repeatable.	#84
	Use <publicationstatus> as an empty element without text or other content with a required attribute @value with limited values and without language attributes. Seeing that the element is part of control and not part of the descriptive section of EAC-CPF, the necessity to have language attribution and an option for internationalisation does not seem to be that prominent. The element is optional and not repeatable.	#85
Topic: Names	No additional agency information in convention declaration as there is no use case available.	#67